

# **Auxin-degron system identifies immediate mechanisms of OCT4**

Lawrence E Bates<sup>\*1</sup>, Mariana R P Alves<sup>2</sup>, José C R Silva<sup>\*134</sup>

<sup>1</sup> Wellcome-MRC Cambridge Stem Cell Institute,  
University of Cambridge,  
Cambridge,  
CB2 0AW,  
UK

<sup>2</sup> Developmental Biology Unit,  
European Molecular Biology Laboratory,  
69117 Heidelberg,  
Germany

<sup>3</sup> Department of Biochemistry,  
University of Cambridge,  
Cambridge,  
CB2 1GA,  
UK

<sup>4</sup> Present address: Centre for Cell Lineage and Atlas,  
Bioland Laboratory,  
Guangzhou Regenerative Medicine and Health Guangdong Laboratory,  
Guangzhou,  
Guangdong,  
510320,  
China

\*Correspondence: lb465@cam.ac.uk, jose\_silva@grmh-gdl.cn

## Abstract

The pluripotency factor OCT4 is essential for the maintenance of naïve pluripotent stem cells in vitro and in vivo. However, the specific role of OCT4 in this process remains unknown. Here, we developed a rapid protein-level OCT4 depletion system that demonstrates that the immediate downstream response to loss of OCT4 is reduced expression of key pluripotency factors. Our data show a requirement for OCT4 for the efficient transcription of several key pluripotency factors, and suggest that expression of trophectoderm markers is a subsequent event. Additionally, we find that NANOG is competent to bind to the genome in the absence of OCT4, and this binding is in fact enhanced. Globally, however, active enhancer associated histone mark H3K27ac is depleted. Our work establishes that while OCT4 is required for the maintenance of the naïve transcription factor network, at a normal ESC level it antagonises this network through inhibition of NANOG binding.

## Introduction

Naïve pluripotent stem cells (nPSCs) are the embryonic founders of the cells present in the adult animal. The transcription factor OCT4, expressed from the gene *Pou5f1*, is necessary for the maintenance of naïve pluripotency in vivo and in vitro. In both cases, loss of OCT4 ultimately leads to exit from naïve pluripotency and with cells taking on characteristics of the trophoblast lineage, including expression of marker genes *Cdx2*, *Pl-1* and *Eomes*, and adoption of trophectoderm-like morphology (Nichols et al., 1998; Niwa et al., 2000).

Despite this, it is not understood why OCT4 is essential to the naïve state; OCT4 is a highly promiscuous transcription factor, binding to thousands of sites in the genome (King and Klose, 2017; Simandi et al., 2016), and many of its target genes are known not to be essential for the maintenance of nPSCs (Hall et al., 2009; Matoba et al., 2006; Zwaka, 2012). At the same time, no target has been identified that can rescue OCT4-loss driven differentiation either through overexpression or knockout (Hall et al., 2009; Matoba et al., 2006). It therefore seems likely that differentiation upon loss of OCT4 is the result of misregulation of a number of genes rather than of any single critical target.

Studying the effects of loss of OCT4 are further complicated by the fact that OCT4 appears to function in a level-dependent manner (Niwa et al., 2000). Mouse nPSCs engineered to express reduced levels of OCT4 are recalcitrant to differentiation and in some cases can be maintained in minimal culture conditions (Karwacki-Neisius et al., 2013; Radzishchanskaya et al., 2013), including basal serum-free media devoid of otherwise essential growth factors and inhibitors of differentiation pathways. Expression of pluripotency associated genes may be enhanced in this state, and responses to signalling cues may be altered. In particular, it appears that transcript and protein levels of the core pluripotency factor *Nanog* are negatively correlated with OCT4 expression (Karwacki-Neisius et al., 2013).

*NANOG* is a homeodomain transcription factor that co-binds with OCT4 at many enhancers to promote expression of other pluripotency genes (Chen et al., 2008; Loh et al., 2006).

While NANOG is not absolutely required for maintenance of the naïve identity (Chambers et al., 2007), loss of NANOG results in widespread differentiation and reduced self-renewal capacity (Chambers et al., 2007; Ivanova et al., 2006; Mitsui et al., 2003). Interestingly, binding of NANOG to the genome appears to be enhanced in self-renewing nPSCs expressing low levels of OCT4, in line with the increased overall level of NANOG protein (Karwacki-Neisius et al., 2013; Radziskeuskaya et al., 2013), while during differentiation induced by loss of OCT4 NANOG binding is reduced (King and Klose, 2017). However, the immediate effect of total loss of OCT4 protein on NANOG binding to the genome remains unknown.

Conventional methods for depleting OCT4, such as genetic ablation via Lox-Cre systems or transcriptional repression using Tet-OFF regulated transgenes, rely on natural degradation to remove Oct4 mRNA and protein. A result of this is that OCT4 persists for a long time. Consequently, responses to depletion of OCT4 are typically examined on the order of hours to days following manipulation (Hall et al., 2009; Niwa et al., 2000), making it impossible to discern whether results are related to OCT4 or simply to cells undergoing differentiation. Further complicating such analysis, microarray data have revealed changes in the expression of thousands of genes at early timepoints after OCT4 suppression, even as the majority of OCT4 protein remains (Hall et al., 2009). This means that such variation may represent a response to a change in OCT4 expression to lower than WT levels, previously described to enhance self-renewal capacity in nPSCs (Karwacki-Neisius et al., 2013; Radziskeuskaya et al., 2013). Thus, the nature of the essential role of OCT4 for the self-renewal of nPSCs remains elusive.

In order to overcome these confounding factors, we generated an OCT4 fusion protein to a full length auxin-inducible degron. This allows us to induce rapid protein-level depletion of OCT4 making it possible for the first time to study the immediate molecular responses to loss of OCT4. These revealed an unprecedented impact on the transcriptional activity of pluripotency-associated transcription factor genes and addressed a long-standing question



regarding the requirement of OCT4 for the binding of a key pluripotency factor, NANOG, to regulatory sequences.

## Results

### **Auxin-degron tagged OCT4 sustains naïve PSC self-renewal and permits rapid loss of OCT4**

OCT4 protein has a relatively long half-life; unlike pluripotency factor NANOG with a reported half-life of around 3 hours, the half-life of OCT4 protein is typically found to be >6 hours to as much as 24 hours (Abranches et al., 2013; Muñoz Descalzo et al., 2013). Additionally, the endogenous *Oct4* mRNA also appears to be unusually stable in mouse ESCs (Abranches et al., 2013). Indeed, we found that conventional tamoxifen-induced CreER driven genetic ablation of *Pou5f1* resulted in a gradual reduction, and required over a day to fully deplete Oct4 RNA and protein (Figure 1A-B). Four days after induction, cells showed distinct morphological changes, resembling trophoblasts and more differentiated trophectodermal cells (Figure S1A). This shows that genetic and transcriptional methods of silencing Oct4 take a long time to fully deplete OCT4 protein, and cells are likely to experience a protracted period of low OCT4 content. Given the transcriptional and phenotypic effects previously observed in cells stably expressing low levels of OCT4 (Karwacki-Neisius et al., 2013; Radziskeuskaya et al., 2013), this is likely to have confounding effects on studies of immediate responses to complete removal of OCT4. We therefore decided to utilize a full length auxin-inducible degron (AID) tag (Baker et al., 2016; Nishimura et al., 2009) to achieve rapid, inducible protein-level degradation of OCT4 on addition of the small molecule IAA.

Since OCT4 is essential for the maintenance of pluripotency we sought to confirm that AID-tagged OCT4 retained its biological function by testing its capacity to support ESCs lacking wild-type OCT4. We utilized an existing TetOFF-Oct4 cell line; non-functional transgenes cannot rescue self-renewal on dox-induced inhibition of wild-type OCT4 in these cells (Niwa

et al., 2002), allowing them to be used in a complementation assay. We generated Oct4<sup>-/-</sup> TetOFF-Oct4 Oct4-AID ESCs (O4AID ESCs hereafter) by transfecting these cells with constitutively expressed Piggybac Oct4-AID and NLS-Tir1 constructs and maintained cells in the presence of doxycycline to inhibit expression of wild-type OCT4 (Figure 1C). Cells continued to proliferate and showed normal morphology (Figure S1B). Cells expressed the fusion protein with no detectable wild-type OCT4 present, and did not show substantially altered expression of key pluripotency genes (Figure 1D-E). This demonstrates that the OCT4 fusion protein retains its original capacity to maintain naïve pluripotency and is not substantially altered in its function by the addition of the AID domain.

To further validate the function of OCT4 within the fusion protein and to establish a second independent cell system, we generated iPSCs null for endogenous OCT4 and therefore wholly reliant on either transgenic wild-type OCT4 or OCT4-AID for their maintenance (Figure 1F). After a period of outgrowth, cells initiated reprogramming and generated intermediate pre-iPSCs (Silva et al., 2008; Theunissen et al., 2011) with high expression of retroviral reprogramming factors but very low expression of the pluripotency factor *Nanog* (Figure 1G). Under naïve-specific conditions, colonies with a domed, iPS-like morphology were readily obtained using either wild-type or Oct4-AID constructs (Figure S1C). *Nanog* was robustly expressed while retroviral transgenes were efficiently silenced in the fully reprogrammed iPSCs (Figure 1G). Again, as expected wild-type OCT4 was not detectable while the OCT4-AID fusion protein was robustly expressed (Figure 1H). Furthermore, in differentiation conditions cells showed loss of pluripotency markers and upregulation of markers of all three germ layers, indicating competency to initiate normal differentiation (Figure S1D). The ability of the Oct4-AID transgene to maintain pluripotency in this OCT4 knockout background further demonstrates that the essential function of OCT4 is maintained.

Utilizing the new protein degradation system, tagged OCT4 protein levels could be greatly reduced within half an hour and undetectable within two hours of addition of IAA (Figure

1E&H). This quick turnover means that there is no appreciable OCT4-low state, and we therefore sought to use this system to study immediate responses of cells to total loss of OCT4.

### **OCT4 is required for the expression of key pluripotency factors**

We analysed gene expression changes following conventional tamoxifen-induced CreER-driven genetic ablation of *Pou5f1* (Figure 2A and Figure S2A). The results indicate that while specific pluripotency-associated genes such as *Nr0b1* are downregulated within 24hrs after induction of OCT4 depletion, upregulation of trophoblast-associated genes occurs concurrently with expression of most naïve pluripotency factors; overall the pluripotency network remains expressed for 24hrs. In keeping with previous observations, however, many of these gene expression changes are detected while significant amounts of OCT4 protein are still present (Figure 1B). This is consistent with reports that expression of pluripotency genes may be enhanced in cells with reduced levels of OCT4 (Karwacki-Neisius et al., 2013; Radzisheuskaya et al., 2013), supporting the notion that conventional methods of OCT4 depletion pass through a protracted OCT4-low state, complicating interpretation of the effects of removal of OCT4.

In contrast to the above, analysis of O4AID cells indicated that expression of naïve-associated factors is quickly extinguished on loss of OCT4 (Figure 2B and Figure S2C-D). Factors such as *Tfcp2l1*, *Klf4* and *Tbx3* are rapidly downregulated following targeted OCT4 protein depletion whereas they were actively maintained in the slower conventional system. Examining several trophoblast-associated genes confirmed that these cells exit pluripotency following loss of OCT4 and differentiate towards trophectodermal lineages (Figure 2B and Figure S2C-D), in keeping with conventional OCT4 depletion systems. However, strong upregulation of trophectoderm associated genes is only seen after a decrease in pluripotency marker expression using this system, indicating that the decision to enter this extraembryonic identity is made only after cells have begun to exit the naïve state. Analysis of eGFPAID ESCs validated that neither loss of pluripotency gene expression nor

upregulation of trophectoderm markers were caused by IAA itself, nor activation of TIR1-induced proteasomal degradation (Figure S2B).

Together these altered kinetics in transcriptional responses highlight that there are important differences between gradually reducing OCT4 protein levels and removing OCT4 protein entirely.

### **Enhancers show rapid epigenetic and functional inactivation following loss of OCT4**

Given the rapid downregulation of pluripotency-associated genes, we examined how enhancer elements were affected by loss of OCT4 at a number of these loci. Enhancer-associated transcription can be observed at many active enhancers (Kim et al., 2010; De Santa et al., 2010), and can play a functional role in promoting transcription of target genes (Alvarez-Dominguez et al., 2017; Shii et al., 2017). We observed loss of transcription from several pluripotency enhancer elements shortly after ablation of OCT4 (Figure 3A), implying loss of enhancer activity.

To validate this, we looked at the level of acetylation of histone H3 at lysine residue 27 (H3K27ac), a mark closely associated with active enhancers. Using ChIP-seq, we observed a dramatic decrease in H3K27ac at pluripotency-associated enhancers (Figure 3B and Figure S3A), and we confirmed this using ChIP-qPCR in both O4AID ES and O4AID iPS systems (Figure S3B-C). Again, we examined eGFPAID ESCs, and found no change in H3K27ac levels at pluripotency-associated enhancers following application of IAA, demonstrating that this response is specifically a response to OCT4 depletion (Figure S3B-C). Loss of H3K27ac was not restricted to pluripotency-associated loci, however; global analysis of all H3K27ac peaks showed a significant decrease (Figure 3C). Using peaks called from OCT4 ChIP-seq data generated by (Marson et al., 2008), we examined if this was the result of decommissioning of OCT4-cobound enhancers; while the reduction in H3K27ac signal was stronger at sites close to OCT4 binding sites, H3K27ac peaks that were not directly bound by OCT4 still showed a marked reduction, indicating a remarkable shift in the global chromatin landscape in the absence of OCT4.

## **OCT4 is dispensable for NANOG binding to pluripotent regulatory sequences**

Depletion of OCT4 was so rapid that protein levels of other pluripotency factors were not strongly affected by the time OCT4 protein was fully removed; NANOG protein levels were only slightly reduced (Figure 4A and Figure S4C). As a result, we decided to use this opportunity to examine changes in the chromatin binding profile of NANOG, a transcription factor that binds to many of the same enhancer elements as OCT4.

NANOG binding efficiency at the enhancer of the *Nanog* and *Pou5f1* genes was analysed by ChIP-qPCR. Surprisingly in both O4AID ES and iPS systems, NANOG not only remained bound but there was greater NANOG signal following depletion of OCT4 (Figure 4B and Figure S4D). Interestingly, other work has suggested that in the OCT4-low state NANOG has increased genomic occupancy, though OCT4 was still present (Karwacki-Neisius et al., 2013; Radziskeuskaya et al., 2013). We validated that Oct4AID was effectively depleted by immunoprecipitating OCT4 from sheared chromatin, followed by Western Blotting. While we could readily pull down Oct4AID from uninduced samples, it was almost undetectable even in the chromatin-enriched eluate in IAA treated samples (Figure S4A). In the data presented here we actively avoid the presence of OCT4, suggesting that either directly or indirectly OCT4 actively reduces the ability of NANOG to bind to enhancer elements.

We therefore examined the NANOG binding profile at the regulatory regions of several key pluripotency genes using ChIP-seq. There was a clear increase in NANOG binding at several loci following depletion of OCT4 (Figure 4C). Interestingly, expression of these genes is decreased following induced OCT4 degradation despite the increased NANOG binding (Figure 2B and Figure S2C-D). Of particular note, the *Klf4* distal enhancer element showed greater NANOG signal following loss of OCT4, yet reduced mRNA and enhancer RNA expression, and H3K27 acetylation (Figure 2B, Figure S2C-D, Figure 3A-B and Figure S3B-C). This further highlights that the presence of OCT4 appears to be absolutely required for the expression of certain pluripotency-associated genes.

### **NANOG binding is increased globally, independent of OCT4 co-binding**

We further analysed our NANOG ChIP-seq data to examine global changes in NANOG binding immediately following OCT4 ablation. Looking at all NANOG binding sites revealed a global shift towards greater levels of NANOG enrichment (Figure 5A and B). Examining the change in NANOG signal over all peaks, there is a significant overall increase following loss of Oct4 (16608 NANOG peaks; paired t-test of NANOG signal before and after treatment,  $p < 10^{-10}$ ; t-test of log fold change of thresholded peaks vs 4458 sub-threshold (background) peaks,  $p < 10^{-10}$ ). To investigate whether the effect on NANOG binding was specific to sites where OCT4 co-binds, we looked at the complementary subsets of NANOG peaks less than 1kb from an OCT4 peak (3930 sites) and those that do not co-bind with OCT4 (12678 sites), again using peaks called from OCT4 ChIP-seq data from (Marson et al., 2008) (Figure 5B). At OCT4-NANOG co-binding sites there was a detectable increase in NANOG signal; globally this was statistically significant (paired t-test of NANOG signal before and after treatment,  $p < 10^{-10}$ ; t-test of log fold change of thresholded peaks vs 382 sub-threshold (background) peaks,  $p < 10^{-10}$ ), though surprisingly the magnitude of the increase was reduced compared to all NANOG peaks. NANOG peaks that do not overlap with OCT4 displayed a more dramatic increase in NANOG binding (12678 peaks; paired t-test of NANOG signal before and after treatment,  $p < 10^{-10}$ ; t-test of log fold change of thresholded peaks vs 4076 sub-threshold (background) peaks,  $p < 10^{-10}$ ), suggesting that the reduced NANOG binding in the presence of OCT4 is not directly due to physical occlusion of NANOG binding sites by OCT4 or due to changes in local chromatin structure in the presence/absence of OCT4. One possible explanation for this global effect is an increase in the stability of NANOG protein following loss of OCT4; however, it has previously shown that NANOG displays increased stability in the presence of OCT4 (Muñoz Descalzo et al., 2013). In keeping with this, we determined the half-life of NANOG with or without addition of IAA which revealed a decrease from ~ 2.3 hours in the presence of OCT4 to ~ 1.4 hours in the absence of OCT4 (Figure S5A). This suggests that stability is not the cause of the increased genomic occupancy of NANOG that we observed.

We additionally examined the effect of chromatin context on the change in NANOG binding following loss of OCT4. We used publicly available ChIP-seq data to classify 2861 NANOG peaks as enhancer- promoter- and/or repressive mark associated. As expected, the majority of these NANOG peaks were associated with enhancers (Figure 5C). In all cases, there was a significant increase in NANOG signal (Figure 5D; paired t-test,  $p < 10^{-10}$ ) and the mean increase in signal was comparable, suggesting that the chromatin context has little impact on the ability of OCT4 to restrict NANOG binding under normal conditions.

## Discussion

It has been shown that OCT4 is dispensable for initial generation of the blastocyst structure, but required for segregation of the inner cell mass into naïve epiblast and primitive endoderm (Le Bin et al., 2014). OCT4 is also essential for maintaining naïve pluripotency in vitro (Niwa et al., 2000). Despite this, the manner in which OCT4 is required for these processes is unclear; no overexpression or knockout mutants that rescue the OCT4 knockout phenotype have been identified (Hall et al., 2009; Matoba et al., 2006; Niwa et al., 2000). In order to infer how OCT4 mediates self-renewal, we wanted to examine initial transcriptional changes in response to OCT4 depletion. It is notable that similar experiments have been performed before (Hall et al., 2009; Matoba et al., 2006) and have failed to yield an essential role for OCT4. However, it is known that cells exhibiting a low level of OCT4 can sustain expression of pluripotency markers under mild differentiation conditions (Karwacki-Neisius et al., 2013; Radziskeuskaya et al., 2013), and we wondered if transcription of such factors was being artificially maintained by the gradually reducing level of OCT4 protein in conventional depletion systems. The half-life of OCT4 is sometimes reported to be very long (Lin et al., 2012; Wei et al., 2007); it took a full day for OCT4 to be fully removed from *Pou5f1<sup>F/-</sup>* cells used in this work (Figure 1A-B), and published work shows more than 10 hours for complete OCT4 depletion using a TetOFF system (Hall et al., 2009; Niwa et al., 2000). This is a significant amount of time, especially since transcriptional changes are already observed within this window. In order to avoid this confounding factor, we

established new cell lines utilising rapid depletion of OCT4 at the protein level (Figure 1C-H). It has previously been reported that the auxin-inducible degron can reduce the half-life of tagged proteins to ~20 minutes in mammalian systems (Holland et al., 2012; Nishimura et al., 2009), and indeed we found that tagged OCT4 protein was greatly reduced within half an hour and fully depleted within 1.5-2 hours (Figure 1E&H). After ensuring that both the AID domain and OCT4 were unaltered in their function in this fusion protein (Figure 1D&G and Figure S1B-C), we proceeded to re-examine the immediate effects of loss of OCT4. Since depletion of OCT4 was so rapid, we were able to examine the effect of loss of OCT4 prior to extensive changes in the protein level of other pluripotency transcription factors. Crucially, the high temporal resolution afforded by such a rapid depletion system reveals two phases of transcriptional change following loss of OCT4, unlike following genetic ablation (Figure 2). First, RNA levels of all pluripotency factors examined were rapidly decreased following degradation of OCT4 protein. Notably, all the factors tested display OCT4 binding at their enhancer or promoter elements in published ChIP-seq data. This suggests that the presence of OCT4 at these elements may be essential for the expression of a broad range of pluripotency associated genes, explaining why no single factor can rescue the OCT4 knockout phenotype. Only subsequently did we observe upregulation of trophectoderm-associated markers, suggesting that this may be a secondary effect or that derepression of these loci occurs over a longer timescale. In keeping with an essential role in permitting active transcription, we found a global reduction in the level of H3K27ac, associated with active enhancers (Figure 3 and Figure S3). The fact that this loss of active enhancer marks extends beyond pluripotency-associated loci could imply that OCT4 is critical for maintaining the uniquely permissive chromatin environment found in naïve pluripotent cells, beyond simply acting to drive expression of key pluripotency factors.

ChIP-qPCR and ChIP-seq against NANOG protein yielded an interesting and surprising result; in the absence of OCT4, greater levels of NANOG were found bound to the genome (Figures 4 and 5, and Figures S4 and 5). Remarkably, this extended beyond OCT4 co-



bound sites, with an increase in NANOG levels observed at many binding sites lacking OCT4 (Figure 5B). Consequently, it seems unlikely that this effect is due to OCT competing with NANOG for binding sites, in keeping with previous reports of simultaneous binding of OCT4 and NANOG to regulatory regions as detected by sequential ChIP (Medeiros et al., 2009). A previous study examining the localisation of NANOG and SOX2 24 hours after transcriptional depletion of OCT4 found that SOX2 and NANOG binding was reduced at many OCT4 co-bound sites following silencing of *Oct4* (King and Klose, 2017). However, this carries the caveat that such changes may be due to subsequent transcriptional and epigenetic changes downstream of loss of OCT4 rather than direct effects. In particular, reduced NANOG signal was associated with loss of transcription and chromatin accessibility, suggesting that these enhancer elements were no longer active or accessible, perhaps due to the onset of differentiation, by the time of their analysis. In contrast, by using a targeted protein degradation system we are able to study the immediate changes in protein localisation in our system, limiting secondary effects.

Interestingly, we observed a rapid downregulation of *Klf4*, an enhancer RNA associated with the *Klf4* super-enhancer, and loss of the active enhancer histone mark H3K27ac at this locus, despite an increase in NANOG binding in this region. It has previously been shown that expression of *Klf4* is highly dependent on the presence of NANOG, although it can be partially rescued with overexpression of ESRRB (Festuccia et al., 2012). However, only when NANOG and STAT3 are both present is there efficient transcription of both *Klf4* and the *Klf4* eRNA; neither overexpression of NANOG in the absence of active STAT3, nor induction of JAK/STAT signalling in NANOG knockout ESCs are capable of inducing *Klf4* or *Klf4* eRNA upregulation (Stuart et al., 2014). Our findings appear to add OCT4 to this list of factors that need to be present to achieve activation of this enhancer region.

It is currently unclear how the global nuclear structure changes following acute depletion of OCT4. Previous work has suggested that OCT4 may act as a pioneer factor, opening chromatin that would otherwise be highly compacted (King and Klose, 2017; Soufi et al.,

2015). As a result, removal of OCT4 might result in the eviction of other transcription factors and repression of gene expression, though the increased binding of NANOG to chromatin suggests this may not be the case (Figures 4 and 5). Additionally, there could be significant changes in the chromatin architecture due to altered topological domains in the absence of OCT4. Decommissioning of super enhancers, many of which are bound by OCT4 in ESCs (Whyte et al., 2013), could lead to a reorganisation of the nuclear structure as these elements interact with multiple promoter regions, even over large distances (Novo et al., 2018).

The ability to probe transcriptional responses immediately following total OCT4 depletion, and the capacity to examine changes in the binding of NANOG to chromatin prior to significant secondary effects has allowed us to go some way to explaining past discrepancies in the perceived role of OCT4 in the pluripotency network. As Hall et al. note, *Nanog* has been described as an OCT4 target gene, and there is significant biochemical data validating that OCT4 binds to the *Nanog* enhancer region and is required for its expression (Kuroda et al., 2005; Rodda et al., 2005), yet gradual OCT4 depletion appeared to have a limited, even positive impact on *Nanog* mRNA (Figure 2A, Figure S2) (Hall et al., 2009). We propose that the reason for this is that while OCT4 may be essential for *Nanog* expression, above a minimal level OCT4 negatively regulates the binding of other transcriptional activators. Thus, as OCT4 levels drop and cells pass through an OCT4-low state, binding of OCT4 to the *Nanog* enhancer is reduced and other transcription factors can be recruited, buffering *Nanog* expression levels. Once OCT4 is entirely depleted *Nanog* expression falls as the locus is no longer able to be transcribed.

In agreement with existing literature, it appears that the half-life of NANOG is increased in the presence of OCT4 (Figure S5A) (Muñoz Descalzo et al., 2013). As a result, increased NANOG stability is not likely to be responsible for the increased genomic occupancy we observed. However, it would be interesting to see what other properties of NANOG protein are altered in the absence of OCT4, particularly whether changes in the NANOG protein

interactome or DNA binding affinity could explain the increase in chromatin binding.

Differences in post-translational modifications could affect the ability of NANOG to interact with DNA and proteins in this way.

Combining the transcriptional responses to loss of OCT4 with the observed global changes in NANOG ChIP signal have led us to a unifying model for the range of phenotypes associated with different levels of OCT4 protein in ESCs (Figure 5E). In the absence of OCT4 key pluripotency factors cannot be expressed and cells cannot maintain a naïve identity. With low levels of OCT4, a threshold is reached such that these factors are able to be expressed, and efficient binding by NANOG results in robust transcription of the whole naïve network to such an extent that differentiation is compromised. At higher OCT4 levels, such as seen in wild-type ESCs and embryos, OCT4 globally suppresses the binding of NANOG resulting in a weaker transcriptional network. This allows cells to undergo differentiation in response to signalling cues. On overexpression of OCT4, it is expected that the binding of NANOG to chromatin will be further reduced, destabilizing the naïve network to the point that it cannot be readily maintained.

In summary, utilizing a rapid protein-level depletion strategy, we identified that the primary transcriptional response to loss of OCT4 is a decrease in the expression of pluripotency-associated genes, and upregulation of trophectoderm factors is a subsequent event. Additionally, we found a global increase in the amount of NANOG associated with the genome in the absence of OCT4, suggesting a mechanism by which wild-type levels of OCT4 ensure that naïve cells retain the capacity to initiate differentiation in response to appropriate signals. Together, these reveal a model that ties together the range of phenotypes associated with differing levels of OCT4, concisely explaining how different levels of this factor result in seemingly contradictory cell behaviours.

## Experimental procedures

### Cell culture

Mouse ESCs were cultured in 2iL conditions as previously described. Briefly, cells were maintained in N2B27 (1:1 DMEM/F-12 and Neurobasal, 2mM L-glutamine, 1x penicillin-streptomycin, 0.1mM 2-mercaptoethanol, 1% B27, 0.5% N2) supplemented with 3μM CHIR99021, 1μM PD0325901 and 20ng/ml mLIF. Where stated cells were transitioned into SL conditions, consisting of GMEM without L-glutamine, 10% fetal bovine serum, 1x non-essential amino acids, 1mM sodium pyruvate, 2mM L-glutamine, 1x penicillin-streptomycin, 0.1mM 2-mercaptoethanol, and 20ng/ml mLIF. ES and iPS cells were maintained on gelatin-coated tissue culture plastic, NSCs were maintained on laminin-coated tissue culture plastic. Cells were passaged every 2-4 days using accutase as required. Where described, media was supplemented with 500nM 4-OHT, 1μg/ml doxycycline, and/or 500μM indole-3-acetic acid.

### Cell Lines

Oct4<sup>F/-</sup> ESCs (Oct4<sup>F/β-geo</sup>) were previously derived from a cross between Oct4<sup>+/-β-geo</sup> and Oct4<sup>F/F</sup> mice. O4AID ESCs were generated by transfecting ZHBTc4 ESCs (Niwa et al., 2000) with pPB-CAG-Oct4AID-PGK-Hph (mouse Oct4 tagged at the C-terminus with full length AID, separated by a short PG linker) and pPB-CAG-Tir1-IRES-Bsd (*Oryza sativa* Tir1, with an N-terminal SV40 NLS signal, codon optimized for expression in mouse) with pPBBase to achieve efficient integration, and maintained in the presence of Doxycycline to silence the TetOFF-Oct4 transgene. O4AID iPSCs were generated from Oct4<sup>-/-</sup> NSCs (Radziszewska et al., 2013) nucleofected with pPB-CAG-Oct4AID-PGK-Hph and pPBBase to achieve efficient integration and reprogrammed; equivalent wild-type Oct4 control cells were generated by nucleofection with pPB-CAG-Oct4-PGK-Hph and pPBBase. The generated O4AID iPSCs were transfected with pCAG-Tir1-IRES-Bsd and clonal lines were assessed for functional transgene integration. eGFP-AID ESCs were generated by transfecting Oct4<sup>F/-</sup> ESCs with

pPB-CAG-eGFPAID-IRES-Zeo (eGFP tagged at the C-terminus with full length AID, separated by a short PG linker) and pPB-CAG-Tir1-IRES-Bsd with pPBase.

### **Reprogramming**

PLAT-E cells were transfected with pMXs-Oct4, pMXs-Klf4 or pMXs-cMyc using FuGENE 6 reagent to produce retroviral particles. Medium was changed the following day and after 48 hours supernatant containing retroviral particles was collected. Media were filtered and mixed together in equal ratio, then polybrene was added to a final concentration of 4µg/ml. The polybrene/virus mixture was applied to NSCs. 24 hours later, NSCs were nucleofected with a 1:5 ration of pPBase and either pPB-CAG-Oct4AID-PGK-Hph or pPB-CAG-Oct4-PGK-Hph using Amaxa Nucleofection Technology and plated in NSC medium for 2 days then switched to SL medium. Medium was then switched to KSR-2iL (GMEM without L-glutamine, 10% KOSR, 1% fetal bovine serum, 1x non-essential amino acids, 1mM sodium pyruvate, 2mM L-glutamine, 1x penicillin-streptomycin, 0.1mM 2-mercaptoethanol, and 20ng/ml mLIF). Selection was applied for expression from the endogenous Oct4 locus on the ninth day in KSR-2iL. Once colonies had expanded they were passaged into 2iL conditions.

### **Conventional Differentiation**

iPSCs were induced to differentiate either in monolayer differentiation (N2B27) or in suspension (embryoid body). For monolayer differentiation, 130k cells were plated in a gelatin-coated well of a 6-well plate in 2iL media. The following day this media was removed, the wells were washed with N2B27 lacking CHIR99021, PD0325901, and LIF, and cells were cultured in N2B27 alone for the duration of the timecourse. For suspension differentiation, 570k cells were transferred to a low-attachment 6cm dish in serum-based medium lacking LIF. Cells were allowed to aggregate and expanded in suspension for the duration of the timecourse. Media was changed by gently centrifuging floating aggregates and transferred back to the low-attachment dish.

### **qRT-PCR and RNA-seq**

Total RNA was isolated from cultured cells using the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions including on-column DNaseI digest. For RT-qPCR, RNA was reverse-transcribed using SuperScript III First-Strand Synthesis SuperMix for RT-qPCR and reactions were performed on a StepOnePlus Real time PCR system with recommending settings using TaqMan Fast Universal PCR Mix or Fast SYBR Green Master Mix. See Supplemental Experimental Procedures for primers and Taqman probes.

For high throughput RNA sequencing, RNA integrity was assessed on a Qubit Fluorometer (ThermoFisher Scientific) and Agilent Bioanalyzer Nano Chips (Agilent Technologies). Depletion of ribosomal RNA was performed on 2-5µg of total RNA using the Ribo-Zero rRNA Removal Kit (Illumina) and libraries were produced from 10-100ng of ribosomal-depleted RNA using NextFlex Rapid Directional RNA-seq Kit (Bio Scientific) with 12 cycles of PCR amplification. Libraries were pooled in equimolar quantities and sequenced on the HiSeq4000 platform (Illumina) at CRUK. Library preparation was performed by the W-MRC CSCI genomics facility. Reads were aligned to mouse genome reference GRCm38/mm10 with TopHat2 (Kim et al., 2013) v2.1.0 (<https://ccb.jhu.edu/software/tophat>) using default parameters for paired end reads. Gene-wise counts were generated using featureCounts (Liao et al., 2014) based on annotation from Ensembl GRCm38.86 release. Transcript counts were TMM normalized and differentially expressed genes were called using a non-parametric algorithm specifically intended for use with data lacking replicates, with  $q > 0.9$  using NOISeq-sim from NOISeq (Tarazona et al., 2011, 2015) version 2.30.0. For all NOISeq analyses, the following parameters were used: pnr=0.2, nss=5, v=0.02; seed '321' was used for each analysis.

### **Western Blot**

Cells were lysed in RIPA buffer (Sigma) containing Complete-ULTRA protease-inhibitor and PhosStop phosphatase-inhibitor cocktails (Roche), and sonicated with Bioruptor200 (Diagenode) at high frequency, alternating 30 s on/off for 3 min. SDS-PAGE electrophoresis

was performed using Bolt 10% Bis-Tris Plus gels (ThermoFisher) in a Novex MiniCell (ThermoFisher). Protein transfer was performed using the semi-dry iBlot2 system (ThermoFisher) and iBlot Transfer Stacks (ThermoFisher). Detection was achieved using HRP-linked secondary antibodies diluted 1:10000 against the appropriate species (GE Healthcare) and ECL Plus Western Blotting Detection System (GE Healthcare). See Supplemental Experimental Procedures for primary antibodies.

### **Protein half-life analysis**

Cells were treated with cycloheximide with or without addition of IAA, and harvested at hourly intervals. Western blot analysis of Nanog protein levels were performed, and a relative standard curve was included using a 1 in 2 dilution series with 4 points from the uninduced sample. Western blots were quantitated using Image Studio Lite Quantification Software.

### **ChIP-qPCR and ChIP-seq**

ChIP was performed as previously described (Radziseuskaya et al., 2013). For detailed protocols see Supplemental Experimental Procedures. Chromatin was analysed by qPCR using a StepOnePlus Real time PCR system with recommending settings and Fast SYBR Green Master Mix. Additionally, next generation sequencing libraries were prepared using the ThruPLEX DNA-Seq Kit. Libraries were pooled in equimolar quantities and sequenced on the HiSeq4000 platform (Illumina) at CRUK. Library preparation was performed by the W-MRC CSCI genomics facility. Reads were aligned to mouse genome reference GRCm38/mm10 using Bowtie 2 (Langmead and Salzberg, 2012) version 2.3.4.3 with default parameters and deduplicated using SAMtools rmdup. For details of next generation sequencing analysis, see Supplemental Experimental Procedures.

### **ChIP-Western Blot**

ChIP was performed as above, except that a single 40 minute elution was performed. Instead of overnight reverse-crosslinking, sample loading buffer and reducing agent were

added to the eluate, and samples were boiled for 40 minutes to denature protein and reverse crosslinks. Western blot was then performed as described above.

## **Data availability**

All sequence data generated during this study are accessible under GEO accession number: GSE168579.

## **Acknowledgments**

We thank S. Lees, M. Paramor, M.A. Barber, and P. Humphreys for specialist technical support. We are grateful to A. Radzisheuskaya for technical assistance and Y. Costa and I. Stockwell for critical reading of the manuscript. This study was supported by start-up funding provided by Bioland Laboratory (Guangzhou Regenerative Medicine and Health Guangdong Laboratory), a Wellcome Trust Fellowship (WT101861) to J.C.R.S and by BBSRC and MRC research grants, BB/R018588/1 and MR/R017735/1 respectively. The authors gratefully acknowledge core support from the Wellcome-MRC Cambridge Stem Cell Institute. This research was funded in whole, or in part, by the Wellcome Trust [203151/Z/16/Z] and the UKRI Medical Research Council [MC\_PC\_17230]. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

## **Author Contributions**

L.E.B. and J.C.R.S. conceived the study. L.E.B. designed and performed experiments and bioinformatics analysis, analysed the data, supervised the study and wrote the manuscript. M.R.P.A. performed experiments. J.C.R.S. supervised the study and wrote and approved the manuscript.

## **Declaration of Interests**

The authors declare no competing interests.



## References

- Abranches, E., Bekman, E., and Henrique, D. (2013). Generation and Characterization of a Novel Mouse Embryonic Stem Cell Line with a Dynamic Reporter of Nanog Expression. *PLoS One* 8, e59928.
- Alvarez-Dominguez, J.R., Knoll, M., Gromatzky, A.A., and Lodish, H.F. (2017). The Super-Enhancer-Derived alncRNA-EC7/Bloodlinc Potentiates Red Blood Cell Development in trans. *Cell Rep.* 19, 2503–2514.
- Baker, O., Gupta, A., Obst, M., Zhang, Y., Anastassiadis, K., Fu, J., and Stewart, A.F. (2016). RAC-tagging: Recombineering And Cas9-assisted targeting for protein tagging and conditional analyses. *Sci. Rep.* 6, 25529.
- Le Bin, G.C., Muñoz-Descalzo, S., Kurowski, A., Leitch, H., Lou, X., Mansfield, W., Etienne-Dumeau, C., Grabole, N., Mulas, C., Niwa, H., et al. (2014). Oct4 is required for lineage priming in the developing inner cell mass of the mouse blastocyst. *Development* 141, 1001–1010.
- Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. *Nature* 450, 1230–1234.
- Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133, 1106–1117.
- Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, S.R., and Chambers, I. (2012). Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. *Cell Stem Cell* 11, 477–490.
- Hall, J., Guo, G., Wray, J., Eyres, I., Nichols, J., Grotewold, L., Morfopoulou, S., Humphreys, P., Mansfield, W., Walker, R., et al. (2009). Oct4 and LIF/Stat3 additively induce Krüppel factors to sustain embryonic stem cell self-renewal. *Cell Stem Cell* 5, 597–609.
- Holland, A.J., Fachinetti, D., Han, J.S., and Cleveland, D.W. (2012). Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 109, E3350–7.
- Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y., and Lemischka, I.R. (2006). Dissecting self-renewal in stem cells with RNA interference. *Nature* 442, 533–538.
- Karwacki-Neisius, V., Göke, J., Osorno, R., Halbritter, F., Ng, J.H., Weiße, A.Y., Wong, F.C.K., Gagliardi, A., Mullin, N.P., Festuccia, N., et al. (2013). Reduced Oct4 expression directs a robust pluripotent state with distinct signaling activity and increased enhancer occupancy by Oct4 and Nanog. *Cell Stem Cell* 12, 531–545.
- Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, R36.
- Kim, T.-K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., et al. (2010). Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465, 182–187.
- King, H.W., and Klose, R.J. (2017). The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells. *Elife* 6.
- Kuroda, T., Tada, M., Kubota, H., Kimura, H., Hatano, S., Suemori, H., Nakatsuji, N., and Tada, T.

- (2005). Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. *Mol. Cell. Biol.* 25, 2475–2485.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.
- Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930.
- Lin, Y., Yang, Y., Li, W., Chen, Q., Li, J., Pan, X., Zhou, L., Liu, C., Chen, C., He, J., et al. (2012). Reciprocal regulation of Akt and Oct4 promotes the self-renewal and survival of embryonal carcinoma cells. *Mol. Cell* 48, 627–640.
- Loh, Y.-H., Wu, Q., Chew, J.-L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., et al. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* 38, 431–440.
- Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J., et al. (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* 134, 521–533.
- Matoba, R., Niwa, H., Masui, S., Ohtsuka, S., Carter, M.G., Sharov, A.A., and Ko, M.S.H. (2006). Dissecting Oct3/4-regulated gene networks in embryonic stem cells by expression profiling. *PLoS One* 1, e26.
- Medeiros, R.B., Papenfuss, K.J., Hoium, B., Coley, K., Jadrach, J., Goh, S.-K., Elayaperumal, A., Herrera, J.E., Resnik, E., and Ni, H.-T. (2009). Novel sequential ChIP and simplified basic ChIP protocols for promoter co-occupancy and target gene identification in human embryonic stem cells. *BMC Biotechnol.* 9, 59.
- Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* 113, 631–642.
- Muñoz Descalzo, S., Rué, P., Faunes, F., Hayward, P., Jakt, L.M., Balayo, T., Garcia-Ojalvo, J., and Martinez Arias, A. (2013). A competitive protein interaction network buffers Oct4-mediated differentiation to promote pluripotency in embryonic stem cells. *Mol. Syst. Biol.* 9, 694.
- Nichols, J., Zevnik, B., Anastasiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* 95, 379–391.
- Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., and Kanemaki, M. (2009). An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat. Methods* 6, 917–922.
- Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.* 24, 372–376.
- Niwa, H., Masui, S., Chambers, I., Smith, A.G., and Miyazaki, J. (2002). Phenotypic complementation establishes requirements for specific POU domain and generic transactivation function of Oct-3/4 in embryonic stem cells. *Mol. Cell. Biol.* 22, 1526–1536.
- Novo, C.L., Javierre, B.-M., Cairns, J., Segonds-Pichon, A., Wingett, S.W., Freire-Pritchett, P., Furlan-Magaril, M., Schoenfelder, S., Fraser, P., and Rugg-Gunn, P.J. (2018). Long-Range Enhancer Interactions Are Prevalent in Mouse Embryonic Stem Cells and Are Reorganized upon Pluripotent State Transition. *Cell Rep.* 22, 2615–2627.
- Radzisheuskaya, A., Chia, G.L. Bin, dos Santos, R.L., Theunissen, T.W., Castro, L.F.C., Nichols, J., and

- Silva, J.C.R. (2013). A defined Oct4 level governs cell state transitions of pluripotency entry and differentiation into all embryonic lineages. *Nat. Cell Biol.* 15, 579–590.
- Rodda, D.J., Chew, J.L., Lim, L.H., Loh, Y.H., Wang, B., Ng, H.H., and Robson, P. (2005). Transcriptional regulation of Nanog by OCT4 and SOX2. *J. Biol. Chem.* 280, 24731–24737.
- De Santa, F., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B.K., Muller, H., Ragoussis, J., Wei, C.-L., and Natoli, G. (2010). A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biol.* 8, e1000384.
- Shii, L., Song, L., Maurer, K., Zhang, Z., and Sullivan, K.E. (2017). SERPINB2 is regulated by dynamic interactions with pause-release proteins and enhancer RNAs. *Mol. Immunol.* 88, 20–31.
- Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., and Smith, A. (2008). Promotion of Reprogramming to Ground State Pluripotency by Signal Inhibition. *PLoS Biol.* 6, e253.
- Simandi, Z., Horvath, A., Wright, L.C., Cuaranta-Monroy, I., De Luca, I., Karolyi, K., Sauer, S., Deleuze, J.-F., Gudas, L.J., Cowley, S.M., et al. (2016). OCT4 Acts as an Integrator of Pluripotency and Signal-Induced Differentiation. *Mol. Cell* 63, 647–661.
- Soufi, A., Garcia, M.F., Jaroszewicz, A., Osman, N., Pellegrini, M., and Zaret, K.S. (2015). Pioneer Transcription Factors Target Partial DNA Motifs on Nucleosomes to Initiate Reprogramming. *Cell*.
- Stuart, H.T., Van Oosten, A.L., Radziszewska, A., Martello, G., Miller, A., Dietmann, S., Nichols, J., and Silva, J.C.R. (2014). NANOG amplifies STAT3 activation and they synergistically induce the naive pluripotent program. *Curr. Biol.* 24, 340–346.
- Tarazona, S., García-Alcalde, F., Dopazo, J., Ferrer, A., and Conesa, A. (2011). Differential expression in RNA-seq: a matter of depth. *Genome Res.* 21, 2213–2223.
- Tarazona, S., Furió-Tarí, P., Turrà, D., Pietro, A. Di, Nueda, M.J., Ferrer, A., and Conesa, A. (2015). Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. *Nucleic Acids Res.* 43, e140.
- Theunissen, T.W., Costa, Y., Radziszewska, A., van Oosten, A.L., Laval, F., Pain, B., Castro, L.F.C., and Silva, J.C.R. (2011). Reprogramming capacity of Nanog is functionally conserved in vertebrates and resides in a unique homeodomain. *Development* 138, 4853–4865.
- Wei, F., Schöler, H.R., and Atchison, M.L. (2007). Sumoylation of Oct4 enhances its stability, DNA binding, and transactivation. *J. Biol. Chem.* 282, 21551–21560.
- Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153, 307–319.
- Zwaka, T.P. (2012). Pluripotency network in embryonic stem cells: maybe Leibniz was right all along. *Cell Stem Cell* 11, 441–442.

## Figure legends

Figure 1. Auxin-degron tagged OCT4 sustains naïve PSC self-renewal and permits rapid loss of OCT4.

A&B) Kinetics of OCT4 depletion in conventional Oct4<sup>F/-</sup> ESCs were examined. A) *Oct4* expression level (RT-qPCR) following addition of 4-OHT and media change to SL. B) OCT4 protein level (Western blot) following addition of 4-OHT and media change to SL.  $\alpha$ -TUBULIN shown as a loading control. C) Schematic showing the generation and use of O4AID ESCs. D) Expression profiling (RT-qPCR) of pluripotency markers. E) Oct4-AID fusion and wild-type OCT4 protein level (Western blot) following addition of IAA.  $\alpha$ -tubulin loading control. F) Schematic showing the generation and use of O4AID iPSCs. rOKM, retroviral Oct4, Klf4 and cMyc. G) Expression profiling (RT-qPCR) of retroviruses and pluripotency factor *Nanog* in partially and fully reprogrammed cells. r.Oct4 – retroviral Oct4; r.cMyc – retroviral cMyc; r.Klf4 – retroviral Klf4. H) Oct4-AID fusion and wild-type protein level (Western blot) following addition of IAA. RT-qPCR data represent mean  $\pm$  SD of three technical replicates. See also Figure S1.

Figure 2. OCT4 is required for the expression of key pluripotency factors

MA plots showing gene expression (RNA-seq) changes following A) deletion of *Pou5f1* in conventional Oct4<sup>F/-</sup> CreER ESCs by addition of 4-OHT or B) degradation of OCT4 protein in O4AID ESCs by addition of IAA. Differentially expressed genes highlighted in green ( $q > 0.9$ , NOISeq-sim). Selected pluripotency (blue) and differentiation (red) associated genes indicated. A) top to bottom: 0 vs 12, 0 vs 18, 0 vs 24, and 0 vs 36 hours respectively. B) top to bottom: 0 vs 1.5, 0 vs 6, 0 vs 10, and 0 vs 24 hours respectively. Note that for Oct4<sup>F/-</sup> CreER ESCs two replicate datasets were merged and for O4AID ESCs a single replicate was used. A non-parametric algorithm designed for use on data lacking replicates was used for the analysis (see Experimental procedures for more details). Key findings were corroborated by RT-qPCR analysis in two independent cell lines, see Figure S2.

Figure 3. OCT4 is required for enhancer activity at key pluripotency loci, and for maintaining global H3K27ac

A) Mapped RNA-seq reads (single replicate) of enhancer RNAs at the *Prdm14* distal enhancer (top) and the *Klf4* distal enhancer (bottom) in O4AID ESCs before and 1.5 hours after addition of IAA. B) ChIP-seq analysis of H3K27ac at the *Prdm14* distal enhancer (top) and the *Klf4* distal enhancer (bottom) in O4AID ESCs before and 1.5 hours after addition of IAA. Genomic coordinates refer to the GRCm38/mm10 genome assembly, and gene intron/exon annotations are taken from Ensembl. OCT4 binding sites generated from ChIP-seq data from Marson et al. 2008 indicated in purple. C) Violin and box plot showing log2-fold change in H3K27ac signal between uninduced and 1.5 hour IAA treated O4AID ESCs. Data were generated by merging mapped ChIP-seq reads from 3 independent IPs. Note that the key findings were corroborated by ChIP-qPCR in two independent cell lines (Figure S3B). Boxes show the median value and extend to the 25<sup>th</sup> and 75<sup>th</sup> quartiles, and whiskers extend to 1.5 times the interquartile range. Plotted are all H3K27ac peaks (n=26457) above a (background) threshold, and the complementary subsets of peaks within 1kb of an OCT4 binding site (n=4662) and not within 1kb of an OCT4 binding site (n=21795). For each set a paired t-test of H3K27ac signal before and after treatment showed highly significant change,  $p < 10^{-10}$ . See also Figure S3.

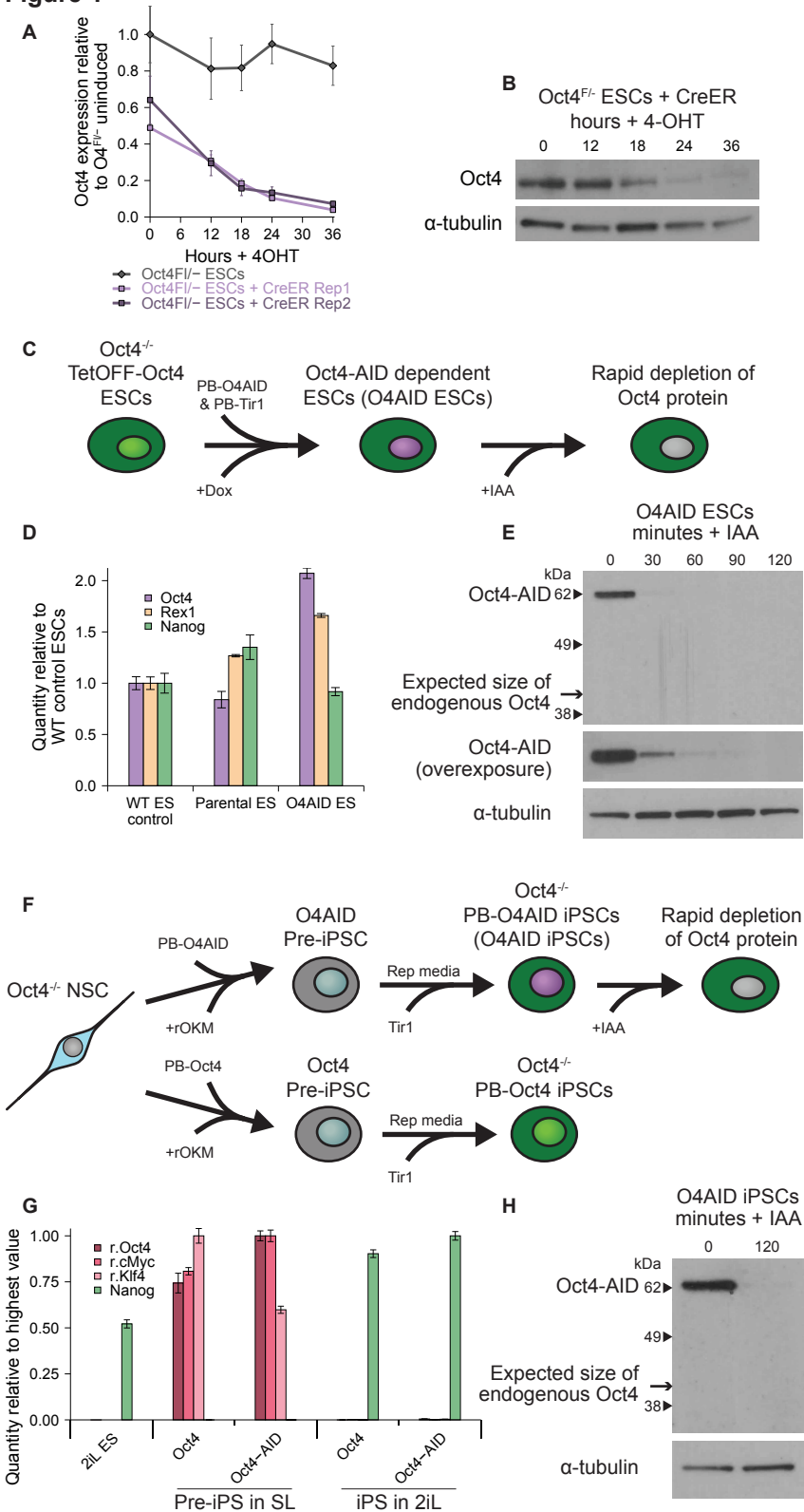
Figure 4. OCT4 is dispensable for NANOG binding to pluripotent regulatory sequences

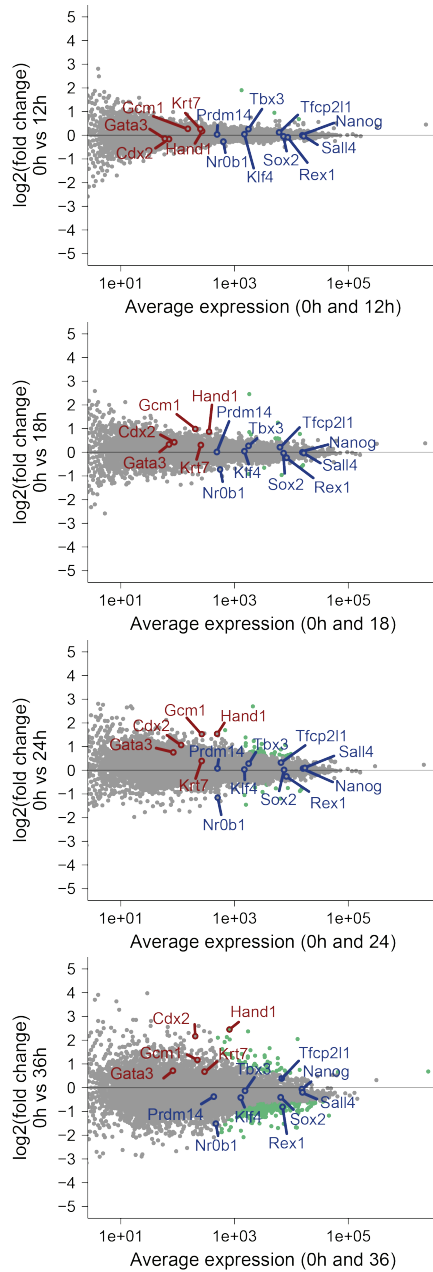
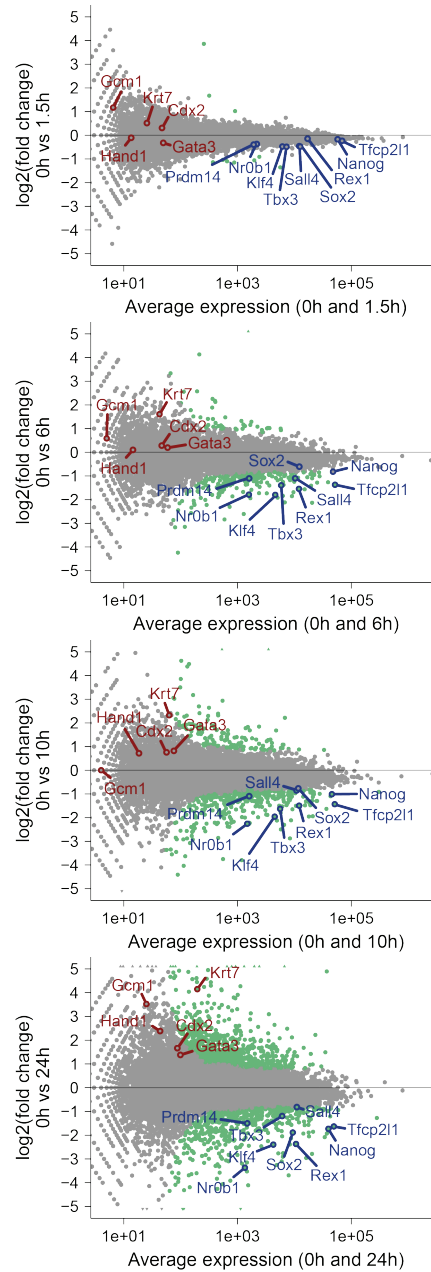
A) Protein level of OCT4 and NANOG (Western blot) in O4AID ESCs before and 1.5 hours after addition of IAA with  $\alpha$ -tubulin as a loading control. B) ChIP qPCR following pulldown of NANOG or using normal IgG negative control at NANOG binding sites or a negative control locus in O4AID ESCs. ChIP qPCR data for NANOG pulldown represent the mean  $\pm$  SD of three IPs; IgG pulldown represents the mean of three technical replicates of single IP. ND = not detected. Note that this is corroborated by ChIP qPCR in an independent cell line (Figure S4). C) Visualization of NANOG ChIP-seq signal (SPKM) across indicated loci before and 1.5 hours after addition of IAA, shown for two IPs. Genomic coordinates refer to the GRCm38/mm10 genome assembly, and gene intron/exon annotations are taken from

Ensembl. OCT4 binding sites generated from ChIP-seq data from Marson et al. 2008 indicated in purple. See also Figure S4.

Figure 5. NANOG binding is increased globally, independent of OCT4 co-binding

A) Summary distribution (top) and heatmap (bottom) of NANOG signal (sum of mapped reads from two independent IPs, normalized to library size) at centred at the summit of all NANOG binding sites across the genome, before and 1.5 hours after addition of IAA. B) Violin and box plot of log2 of the fold change in the average normalized NANOG signal at each NANOG peak in the genome (n=16608) before and after addition of IAA, further broken down into loci within (n=3930) or further than (n=12678) 1kb of OCT4 binding sites generated from ChIP-seq data from Marson et al. 2008. Mapped reads from two independent IPs each before and after treatment were merged to generate NANOG ChIP-seq data. Boxes show the median value and extend to the 25<sup>th</sup> and 75<sup>th</sup> quartiles, and whiskers extend to 1.5 times the interquartile range. C) Euler plot showing the number of NANOG peaks assigned to various chromatin environments, and the overlap between assignments. Numbers indicate the number of peaks uniquely in that section of the diagram. D) Violin and box plot of log2 of the fold change in the average normalized NANOG signal (data as in B) at each NANOG peak in the genome before and 1.5 hours after addition of IAA, further broken down into non-exclusive chromatin environments, as indicated in C. Boxes show the median value and extend to the 25<sup>th</sup> and 75<sup>th</sup> quartiles, and whiskers extend to 1.5 times the interquartile range. E) Model indicating the proposed relationship between the quantity of OCT4 protein and the capacity for NANOG to bind to the genome and consequently the ability of cells to maintain a naïve identity. See also Figure S5.

**Figure 1**

**Figure 2****A** Gene expression changes  
in Oct4<sup>tr</sup> ESCs + 4-OHT**B** Gene expression changes  
in O4AID ESCs + IAA

● Differentially expressed gene  
● Pluripotency-associated gene  
● Trophectoderm-associated gene



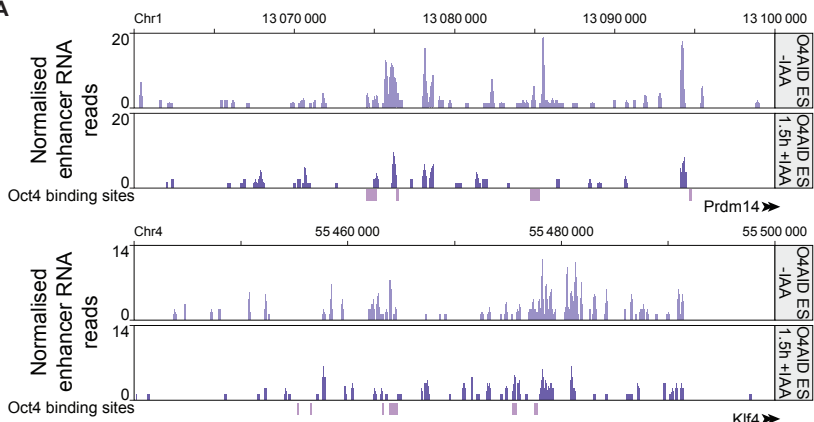
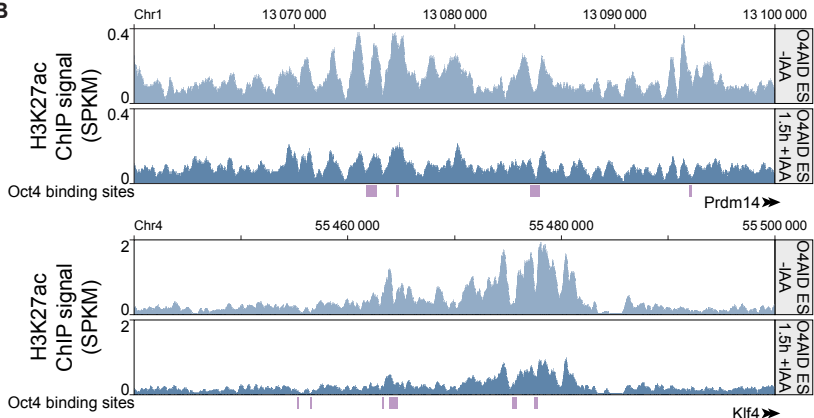
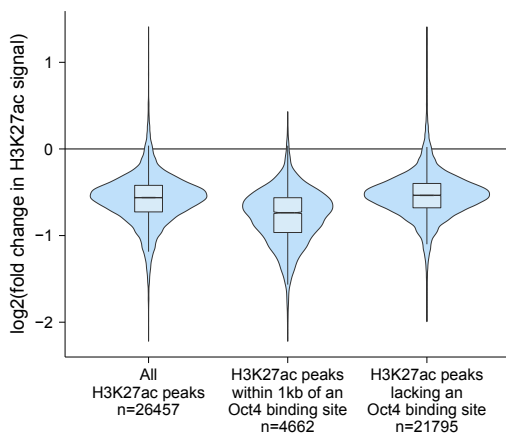
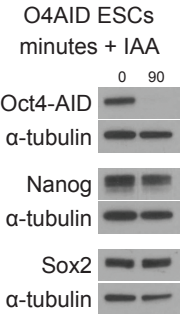
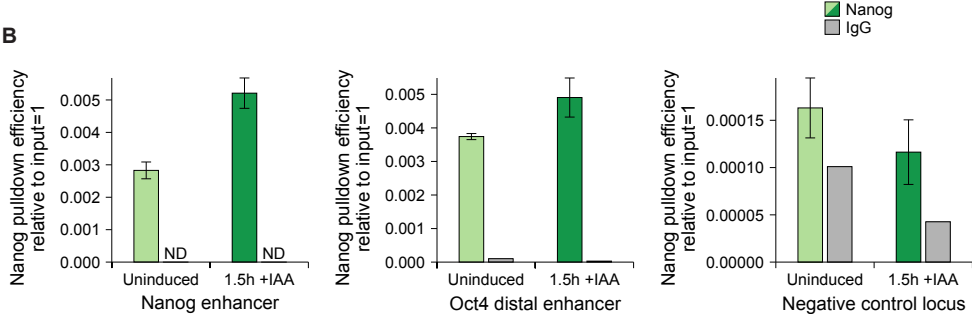
**Figure 3****A****B****C**

Figure 4

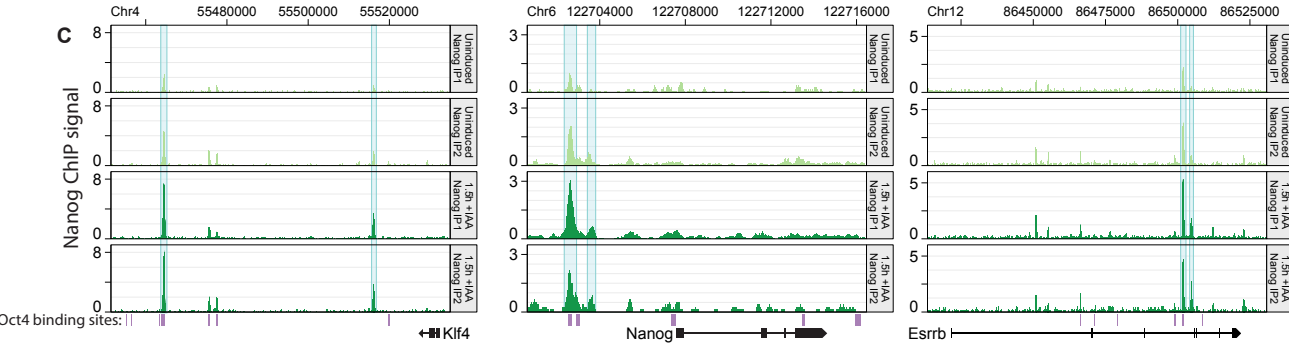
A



B

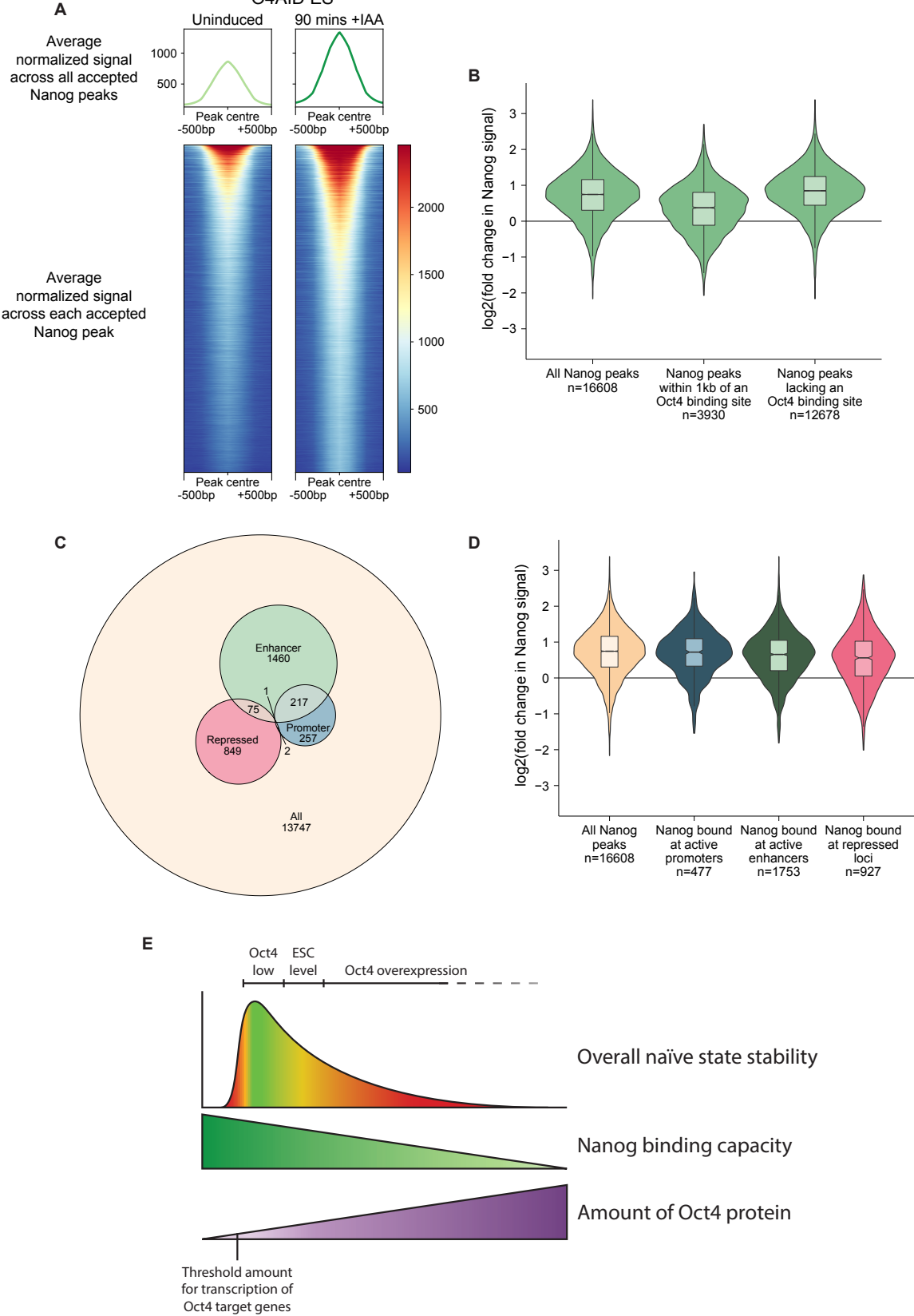


C



**Figure 5**

**Nanog ChIP  
O4AID ES**



**Figure S1**

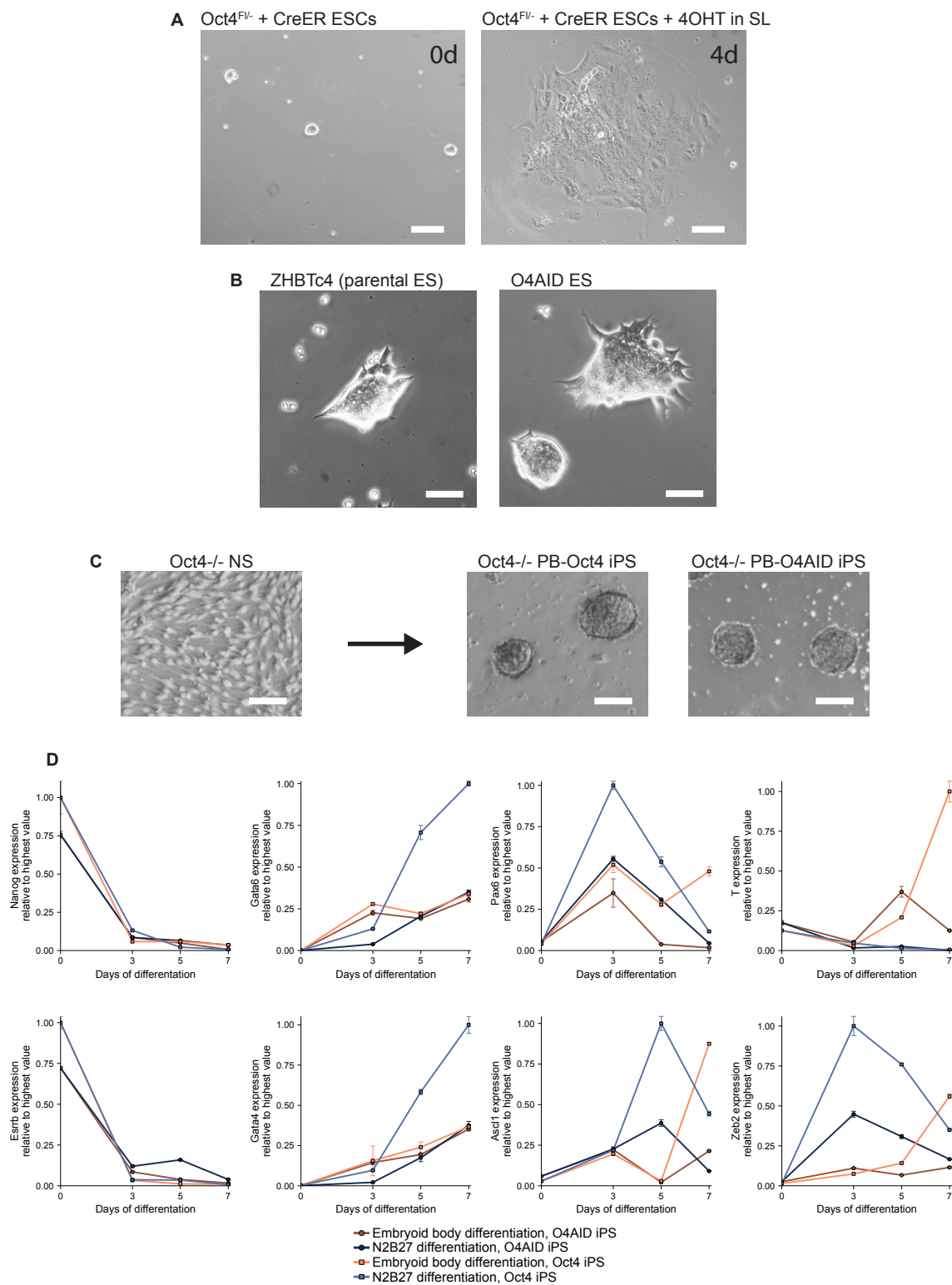


Figure S1. Characterization of cell lines. Relating to Figure 1.

A) Cells before and after conventional tamoxifen-induced CreER driven genetic ablation of Oct4. Uninduced cells grow as compact, round, domed colonies. 4 days after application of tamoxifen, cells have attained an enlarged, flattened morphology. Scalebar – 100 $\mu$ m. B) Oct4<sup>-/-</sup> Tet-OFF Oct4 ESCs (left) and O4AID ESCs share similar morphology, with mostly compact, round, domed colonies. Scalebar – 50 $\mu$ m. C) Oct4<sup>-/-</sup> NSCs (left) were reprogrammed in the presence of either wild-type Oct4 or Oct4-AID fusion protein. In both cases, similar round, domed colonies were formed (right). Scalebar – 100 $\mu$ m. D) RT-qPCR analysis of pluripotency and lineage markers following monolayer differentiation (N2B27) and suspension differentiation (embryoid body) of O4AID iPS and Oct4 iPSCs. RT-qPCR data represent the mean  $\pm$  SD of three technical replicates.

Figure S2

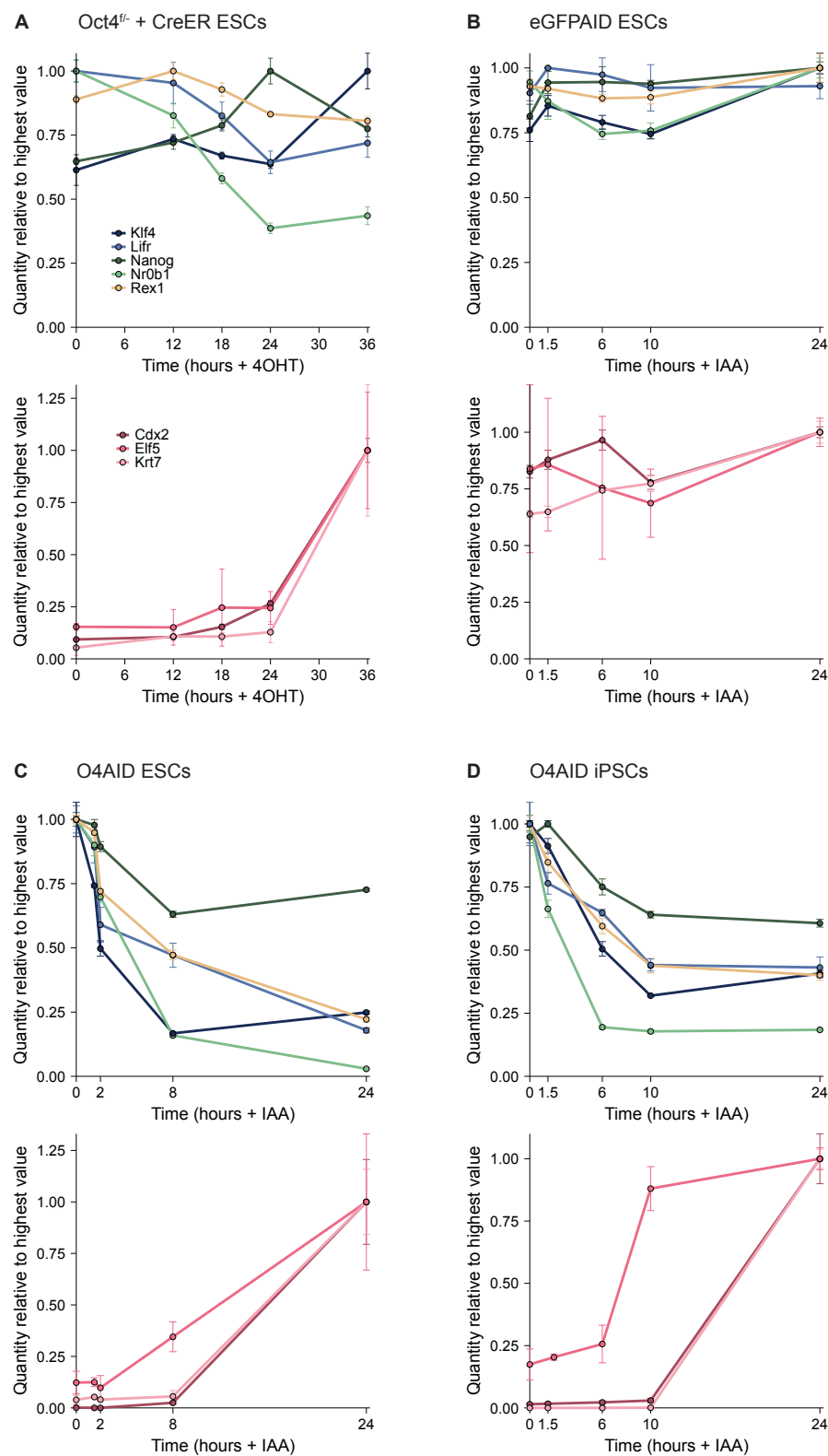


Figure S2. Expression changes following induction. Relating to Figure 2.

RT-qPCR analysis of pluripotency (top) and trophectoderm (bottom) associated genes in Oct4F/- CreER ESCs (A), eGFP AID ESCs (B), O4AID ESCs (C), and O4AID iPSCs (D), in a timecourse following induced depletion of Oct4 or eGFP with 4OHT or IAA. RT-qPCR data represent the mean  $\pm$  SD of three technical replicates.

Figure S3

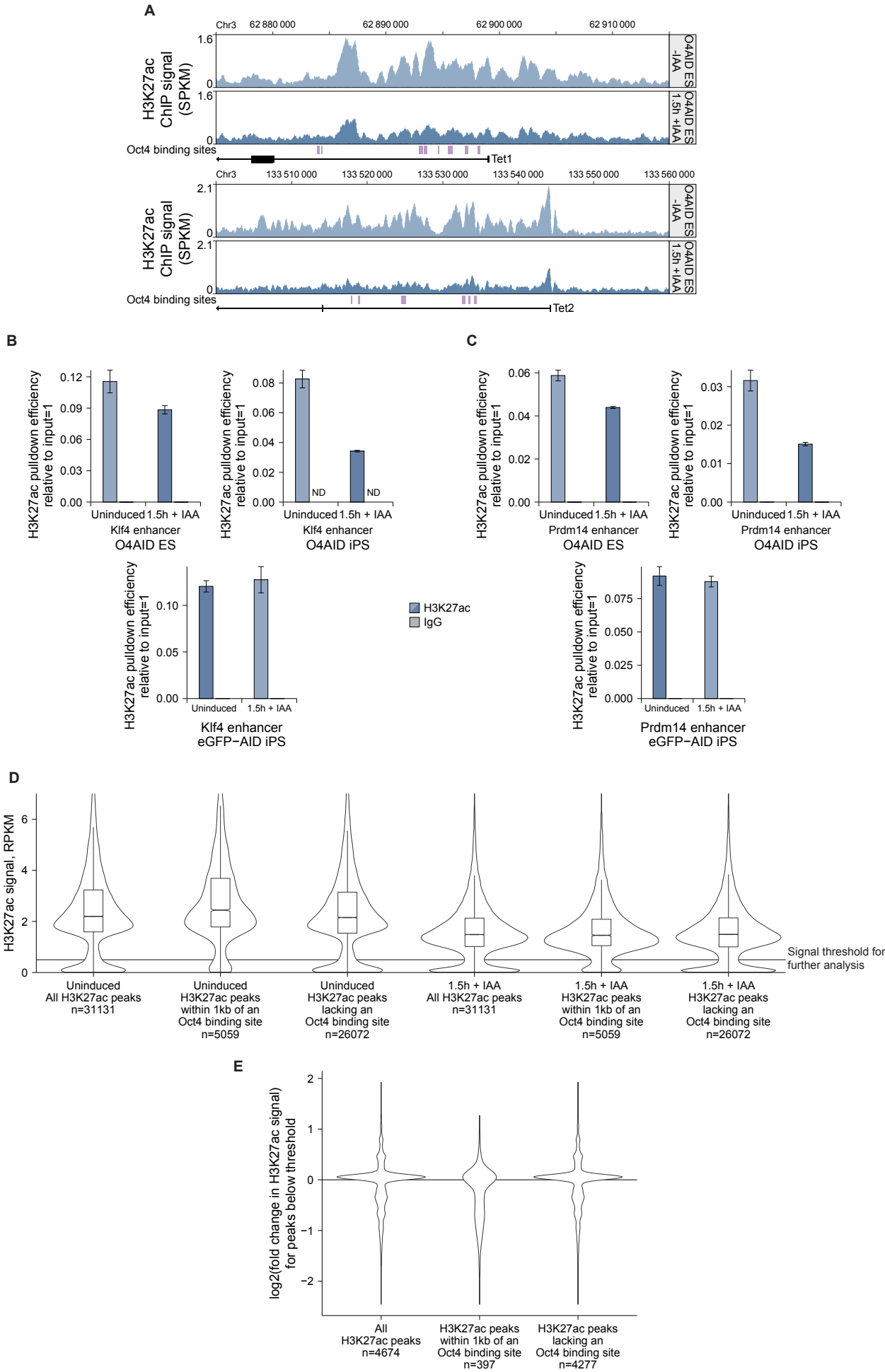




Figure S3. H3K27ac ChIP analysis following induction. Relating to Figure 3.

A) ChIP-seq analysis of H3K27ac at the Tet1 enhancer (top) and the Tet2 enhancer (bottom) in O4AID ESCs before and 1.5 hours after addition of IAA. Genomic coordinates refer to the GRCm38/mm10 genome assembly, and gene intron/exon annotations are taken from Ensembl. Oct4 binding sites generated from ChIP-seq data from Marson et al. 2008 indicated in purple. B&C) ChIP-qPCR analysis of H3K27ac at B) the Klf4 distal enhancer and C) the Prdm14 distal enhancer in O4AID ESCs (left), O4AID iPSCs (right) and eGFPAID ESCs (bottom) before and 1.5 hours after addition of IAA. ChIP-qPCR data represent the mean  $\pm$  SD of three IPs for H3K27ac, and the mean of three technical replicates of a single IP for IgG. ND = not detected. D) Violin and box plot showing H3K27ac signal before and 1.5 hours after addition of IAA at all detected H3K27ac peaks. Peaks with signal below the indicated threshold both before and after induction were considered to be background and were excluded from further analysis. E) Violin plot showing log2-fold change in H3K27ac signal between uninduced and 1.5 hour IAA treated O4AID ESCs at the peaks below the threshold indicated in D).

Figure S4

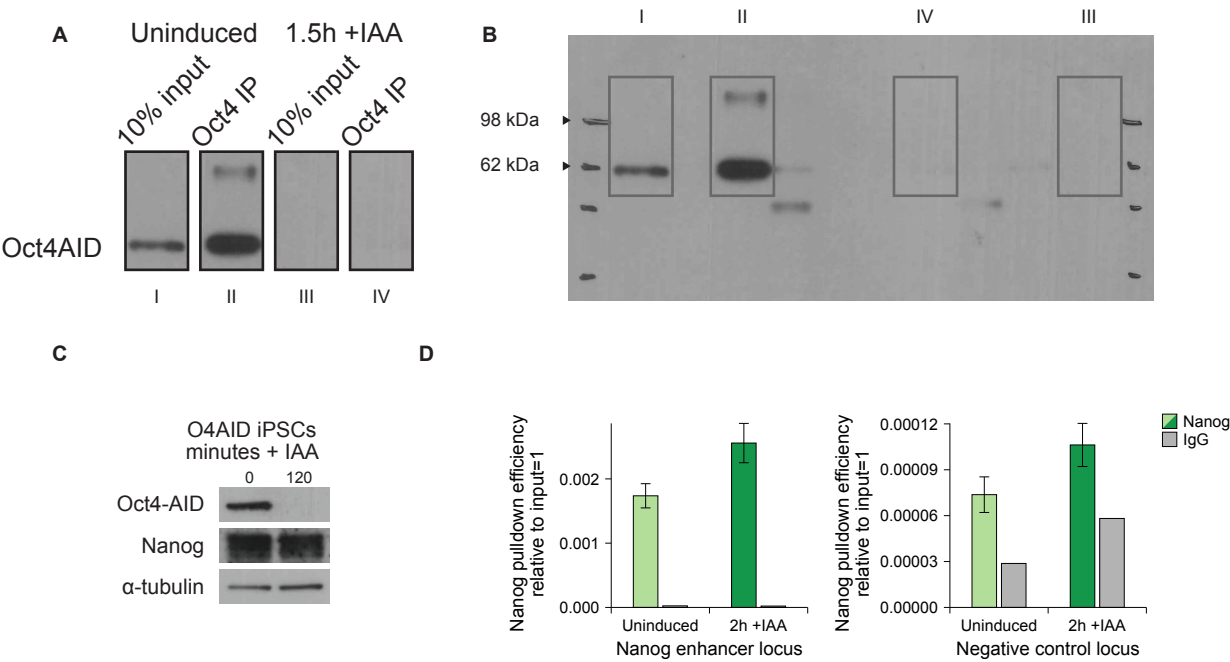


Figure S4. Analysis of Oct4 and Nanog following induction. Relating to Figure 4.

A) ChIP-Western for Oct4AID in O4AID ESCs, either uninduced or following 1.5 hours of treatment with IAA. Input samples were taken from sheared, crosslinked chromatin prior to immunoprecipitation. Input was diluted so that one tenth the amount was loaded compared to IP samples. Note, a high molecular weight band in the uninduced pull-down lane likely represents Oct4 still bound to chromatin due to incomplete reverse crosslinking. Non-neighbouring lanes from the same blot have been aligned for comparison (full blot in B)). B) Complete blot that lanes in A) are taken from. C-D) Analysis of Nanog binding to pluripotent regulatory sequences in O4AID iPSCs. C) Protein level of Oct4 and Nanog (Western blot) in O4AID iPSCs before and 2 hours after addition of IAA with  $\alpha$ -tubulin as a loading control. D) ChIP qPCR following pulldown of Nanog or using normal IgG negative control at Nanog binding sites or a negative control locus in O4AID iPSCs. Nanog pulldown represents mean  $\pm$  SD of three IPs; IgG pulldown represents the mean of three technical replicates of a single IP.

Figure S5

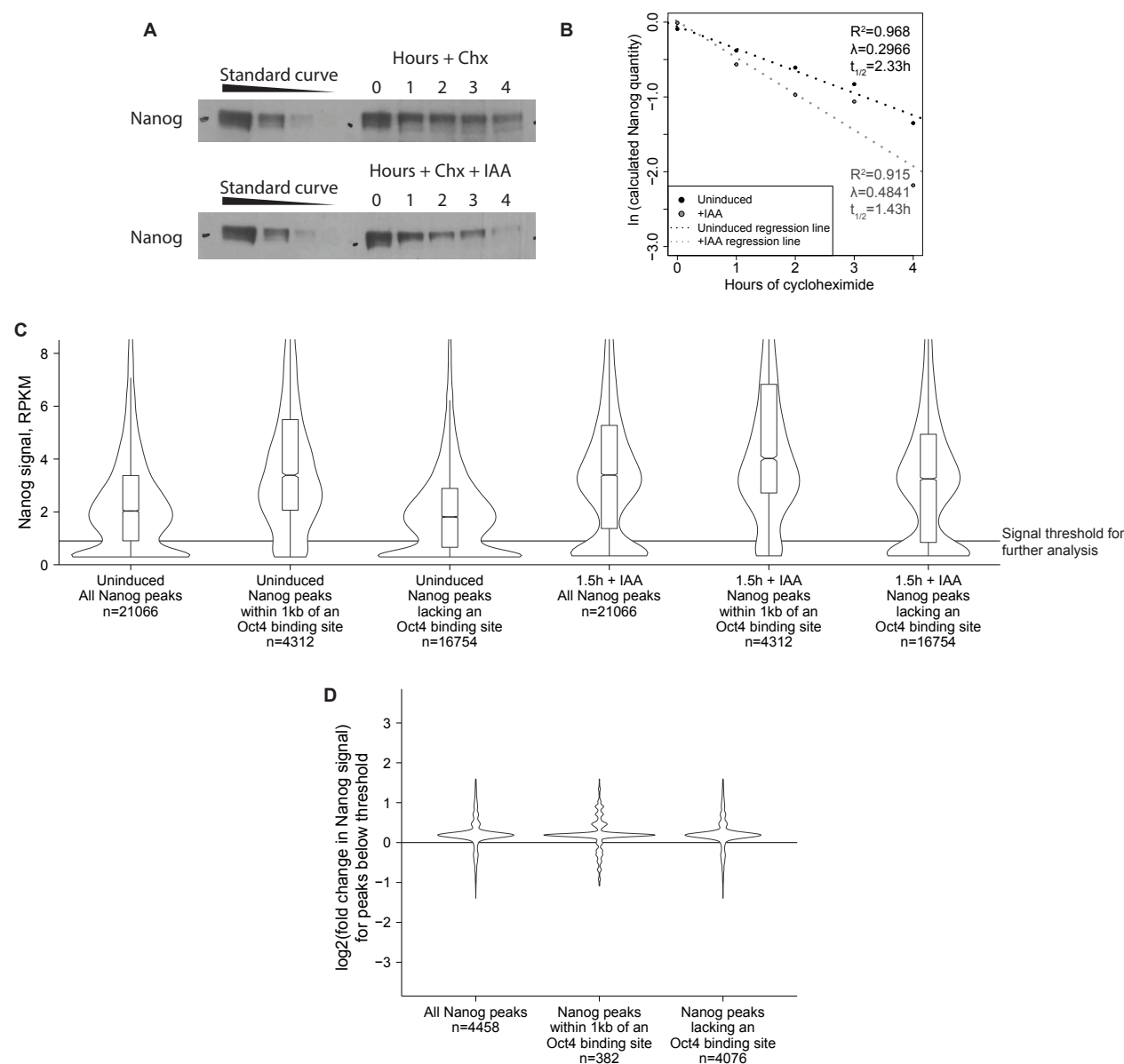


Figure S5. Analysis of Nanog stability and global chromatin binding. Relating to Figure 5.

A&B) Analysis of Nanog protein half-life in O4AID ES with and without Oct4 depletion. A) Western blot analysis of Nanog protein level in a timecourse following cycloheximide treatment, with or without addition of IAA. Standard curve is equivalent to 1X 0h uninduced sample, 0.5X, 0.25X, 0.125X. B) Plot of natural log of Nanog protein quantity against time used to calculate the decay constant,  $\lambda$ , from the equation  $\ln(x(t)) = \ln(x_0) + \lambda t$ . From this, the protein half-life is calculated as  $t_{1/2} = \ln(2)/\lambda$ . C) Violin and box plot showing Nanog signal before and 1.5 hours after addition of IAA at all detected Nanog peaks. Peaks with signal below the indicated threshold both before and after induction were considered to be background and were excluded from further analysis. D) Violin plot showing log2-fold change in Nanog signal between uninduced and 1.5 hour IAA treated O4AID ESCs at the peaks below the threshold indicated in C).

## Supplemental Experimental Procedures

### Chromatin immunoprecipitation

Briefly,  $10 \times 10^6$  cells per IP for transcription factors or  $5 \times 10^6$  cells per IP for histone modifications were fixed for 10 min in 1% formaldehyde. Fixation was halted by addition of an excess of glycine. Cells were washed with ice-cold PBS. Nuclei were isolated: Cells were incubated with Lysis Buffer 1 (50 mM HEPES pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% Glycerol, 0.5% NP40, 0.25% Tx100) at 4 °C for 10 mins, pelleted, then incubated with Lysis Buffer 2 (10 mM Tris pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA) at 4 °C for 10 mins. Nuclei were pelleted then resuspended in Shearing Buffer (50 mM Tris pH 8.0, 1% SDS, 10 mM EDTA) and sonicated to obtain an average DNA fragment size of 500 base pairs. Lysates were diluted 1:10 in dilution buffer (50 mM Tris-HCl at pH 8.0, 167 mM NaCl, 1.1% Triton X-100 and 0.11% sodium deoxycholate) and pre-cleared for 2 h at 4 °C with Dynabeads Protein G magnetic beads (Invitrogen) that were pre-incubated with isotypic IgG antibody. Supernatant was collected, a 10% input sample was taken for relative quantitation, and the chromatin was incubated overnight at 4 °C with 2 µg of antibody or 2 µg of isotypic IgG control. Lysates were then incubated for 1 h at 4 °C with BSA-blocked Dynabeads Protein G magnetic beads, and the beads were washed twice in wash buffer 1 (50 mM Tris pH 8.0, 0.1% SDS, 0.1% Sodium deoxycholate, 1% Tx100, 150 mM NaCl, 1 mM EDTA, 0.5 mM EGTA), once in Wash Buffer 2 (50 mM Tris pH 8.0, 0.1% SDS, 0.1% sodium deoxycholate, 1% Tx100, 500 mM NaCl, 1 mM EDTA, 0.5 mM EGTA), once in Wash Buffer 3 (50 mM Tris pH 8.0, 250 mM LiCl, 0.5% Sodium deoxycholate, 0.5% NP40, 1 mM EDTA, 0.5 mM EGTA) and twice in wash buffer 4 (50 mM Tris pH 8.0, 10 mM EDTA and 5 mM EGTA). Chromatin was eluted twice for a total of 30 min at 65 °C in elution buffer (1% SDS and 0.1 M NaHCO<sub>3</sub>). NaCl was added to samples and inputs to a final concentration of 200 mM, before incubating overnight at 65 °C to reverse the crosslinking and DNA was purified using the QIAquick PCR Purification kit (Qiagen).

### ChIP-seq analysis

Normalized enrichment tracks were generated from aligned reads with the MACS2 (Liu, 2014) callpeak function using the `-SPMR` flag. H3K27ac peaks were called using MACS2 callpeak function on pooled mapped reads of 3 independent IPs from uninduced samples. For analysis of H3K27ac signal intensity, mapped reads from 3 independent IPs were pooled and RPKM normalized using DiffBind (Stark and Brown, 2017). Peaks with RPKM < 0.5 in both uninduced and induced samples were considered to be background and excluded from further analysis. Subsets of peaks located within 1kb of Oct4 peaks were determined by analysis of published ChIP-seq data; reads were deduplicated and aligned as above, and peaks were called using MACS2 callpeak function with whole cell extract used as control samples.

Nanog peaks were called from aligned reads using MACS2 callpeak function on pooled ChIP-seq data from two independent IPs each from both uninduced and IAA treated samples. For analysis of Nanog signal at binding sites, mapped reads from two independent IPs each of treated and untreated samples were pooled and RPKM normalized using deepTools2 (Ramírez et al., 2016) bamCoverage and blacklisted regions from the ENCODE project consortium were excluded, then a heatmap and summary plot were generated using deepTools2 computeMatrix and plotHeatmap tools. For pairwise analysis of Nanog binding at peaks called from pooled data from both uninduced and IAA treated samples, centred on the peak summit and extending a total of 350bp, replicates were pooled and RPKM normalized using DiffBind. Peaks with an RPKM < 0.9 in both uninduced and induced samples were considered to be background and excluded from further analysis. Subsets of peaks corresponding to active promoters, active enhancers and repressed regions were assigned using published ChIP-seq peaksets from ENSEMBL, using the following criteria: active promoters, within 1kb of PolII and either overlapping H3K4me1 or H3K4me3; active enhancer, within 1kb of P300 and overlapping H3K27ac; repressed region, overlapping either H3K9me3 or H3K27me3.

### Public datasets used

Dataset	Source	Analysis
Bruce4 ES EP300 ChIP-seq peakset	Ensembl (release 100)	SWEmbl

Bruce4 ES PolII ChIP-seq peakset	Ensembl (release 100)	SWEmbl
Bruce4 ES H3K27ac ChIP-seq peakset	Ensembl (release 100)	SWEmbl
Bruce4 ES H3K27me3 ChIP-seq peakset	Ensembl (release 100)	CCAT
Bruce4 ES H3K4me1 ChIP-seq peakset	Ensembl (release 100)	CCAT
Bruce4 ES H3K4me3 ChIP-seq peakset	Ensembl (release 100)	SWEmbl
Bruce4 ES H3K9me3 ChIP-seq peakset	Ensembl (release 100)	CCAT
V6.5 ES Oct4 ChIP-seq data	Marson et al. 2008, GEO accession GSM307137	See above
V6.5 ES whole cell extract data	Marson et al. 2008, GEO accessions GSM307154 and GSM307155	See above

## Antibodies

Target	Source	Catalogue number
Polyclonal goat anti-Oct4	Santa Cruz Biotechnology	sc-8628
Monoclonal mouse anti-alpha-Tubulin	Abcam	ab7291
Polyclonal rabbit anti-H3K27ac	Abcam	ab4729
Monoclonal rat anti-Nanog	eBioscience	14-5761-80
Monoclonal rat anti-Sox2	eBioscience	14-9811-80

## qPCR primers and probesets

Target	Use	Source	Catalogue number / sequence
Oct4	RT-qPCR	ThermoFisher Scientific	Mm00658129_gH
Nanog	RT-qPCR	ThermoFisher Scientific	Mm02384862_g1
Rex1	RT-qPCR	ThermoFisher Scientific	Mm03053975_g1
Klf4	RT-qPCR	ThermoFisher Scientific	Mm00516104_m1
Lifr	RT-qPCR	ThermoFisher Scientific	Mm00442942_m1
Nr0b1	RT-qPCR	ThermoFisher Scientific	Mm00431729_m1
Cdx2	RT-qPCR	ThermoFisher Scientific	Mm01212280_m1
Elf5	RT-qPCR	ThermoFisher Scientific	Mm00468732_m1
Krt7	RT-qPCR	Sigma Aldrich	MMUS_NM_033073_1
Gata6	RT-qPCR	ThermoFisher Scientific	Mm00802636_m1
Pax6	RT-qPCR	ThermoFisher Scientific	Mm00443081_m1
T	RT-qPCR	ThermoFisher Scientific	Mm01318252_m1
Gata4	RT-qPCR	ThermoFisher Scientific	Mm00484689_m1
Ascl1	RT-qPCR	ThermoFisher Scientific	Mm03058063_m1
Zeb2	RT-qPCR	ThermoFisher Scientific	Mm00497193_m1
Gapdh	RT-qPCR	ThermoFisher Scientific	4352339E
Retroviral Oct4	RT-qPCR	Sigma Aldrich	F: TGGTACGGGAAATCACAAGTTTGT R: GGTGAGAAGGCGAAGTCTGAAG

			Probe: FAM-CACCTTCCCCATGGCTG-MGB
Retroviral cMyc	RT-qPCR	Sigma Aldrich	F: TGGTACGGGAAATCACAAGTTTGT R: GGTCATAGTTCCTGTTGGTGAAGTT Probe: FAM-CCCTTCACCATGCCCC-MGB
Retroviral Klf4	RT-qPCR	Sigma Aldrich	F: TGGTACGGGAAATCACAAGTTTGT R: GAGCAGAGCGTCGCTGA Probe: FAM-CCCCTTCACCATGGCTG-MGB
Klf4 enhancer	ChIP-qPCR	Sigma Aldrich	F: TGTCTCTCCACTCCCACAA R: AGGAGTGACTGCGTCAAACA
Prdm14 enhancer	ChIP-qPCR	Sigma Aldrich	F: CTACCCTAGACCCACCTCC R: GGGTCCCCTCTCAAAACACC
Nanog enhancer	ChIP-qPCR	Sigma Aldrich	F: CCCACCTGTCCCTAGTCCCCGCT R: TTGGAAGTCTGTGTGGGTGGGG
Oct4 distal enhancer	ChIP-qPCR	Sigma Aldrich	F: GCATAACAAAGGTGCATGATAGCT R: AAATAAAGGCAGCGACTTGGAA
Chr6 negative control locus	ChIP-qPCR	Sigma Aldrich	F: ACTACCCAACTATTGCTCCTGA R: GCTTAACCTGCTCTCCAGG

### Supplemental References

Liu, T. (2014). Use model-based Analysis of ChIP-Seq (MACS) to analyze short reads generated by sequencing protein-DNA interactions in embryonic stem cells. *Methods Mol. Biol.* 1150, 81–95.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J., et al. (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* 134, 521–533.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 44, W160-5.

Stark, R., and Brown, G. (2017). DiffBind: differential binding analysis of ChIP-seq peak data. <http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf>